



OIPE

RAW SEQUENCE LISTING

DATE: 01/17/2002

PATENT APPLICATION: US/10/024,933

TIME: 17:50:40

Input Set : N:\Crif3\RULE60\10024933.raw

Output Set: N:\CRF3\01172002\J024933.raw

SEQUENCE LISTING

1 (1) GENERAL INFORMATION:

2 (i) APPLICANT: Bandman, Olga

3 Lal, Preeti

4 Corley, Neil C.

5 Shah, Purvi

6 (ii) TITLE OF INVENTION: HUMAN S-ANDENOSYL-L-METHIONINE

7 METHYLTRANSFERASE

8 (iii) NUMBER OF SEQUENCES: 4

9 (iv) CORRESPONDENCE ADDRESS:

10 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

11 (B) STREET: 3174 Porter Drive

12 (C) CITY: Palo Alto

13 (D) STATE: CA

14 (E) COUNTRY: USA

15 (F) ZIP: 94304

16 (v) COMPUTER READABLE FORM:

17 (A) MEDIUM TYPE: Diskette

18 (B) COMPUTER: IBM Compatible

19 (C) OPERATING SYSTEM: DOS

20 (D) SOFTWARE: FastSEQ for Windows Version 2.0

21 (vi) CURRENT APPLICATION DATA:

C--> 22 (A) APPLICATION NUMBER: US/10/024,933

C--> 23 (B) FILING DATE: 18-Dec-2001

24 (C) CLASSIFICATION:

25 (vii) PRIOR APPLICATION DATA:

26 (A) APPLICATION NUMBER: 09/149,534

27 (B) FILING DATE:

28 (viii) ATTORNEY/AGENT INFORMATION:

29 (A) NAME: Billings, Lucy J.

30 (B) REGISTRATION NUMBER: 36,749

31 (C) REFERENCE/DOCKET NUMBER: PF-0352 US

32 (ix) TELECOMMUNICATION INFORMATION:

33 (A) TELEPHONE: 415-855-0555

34 (B) TELEFAX: 415-845-4166

35 (C) TELEX:

36 (2) INFORMATION FOR SEQ ID NO: 1:

37 (i) SEQUENCE CHARACTERISTICS:

38 (A) LENGTH: 281 amino acids

39 (B) TYPE: amino acid

40 (C) STRANDEDNESS: single

41 (D) TOPOLOGY: linear

42 (vii) IMMEDIATE SOURCE:

RAW SEQUENCE LISTING
 PATENT APPLICATION. US/10/024,933

DATE: 01/17/2002
 TIME: 17:50:40

Input Set : N:\Crif3\RULE60\10024933.raw
 Output Set: N:\CRF3\01172002\J024933.raw

```

43      (A) LIBRARY: THP1PLB01
44      (B) CLONE: 10625
45      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
46      Met Ala Ser Arg Gly Arg Arg Pro Glu His Gly Gly Pro Pro Glu Leu
47      1          5          10          15
48      Phe Tyr Asp Glu Thr Glu Ala Arg Lys Tyr Val Arg Asn Ser Arg Met
49      20          25          30
50      Ile Asp Ile Gln Thr Arg Met Ala Gly Arg Ala Leu Glu Leu Leu Tyr
51      35          40          45
52      Leu Pro Glu Asn Lys Pro Cys Tyr Leu Leu Asp Ile Gly Cys Gly Thr
53      50          55          60
54      Gly Leu Ser Gly Ser Tyr Leu Ser Asp Glu Gly His Tyr Trp Val Gly
55      65          70          75          80
56      Leu Asp Ile Ser Pro Ala Met Leu Asp Glu Ala Val Asp Arg Glu Ile
57      85          90          95
58      Glu Gly Asp Leu Leu Leu Gly Asp Met Gly Gln Gly Ile Pro Phe Lys
59      100         105         110
60      Pro Gly Thr Phe Asp Gly Cys Ile Ser Ile Ser Ala Val Gln Trp Leu
61      115         120         125
62      Cys Asn Ala Asn Lys Lys Ser Glu Asn Pro Ala Lys Arg Leu Tyr Cys
63      130         135         140
64      Phe Phe Ala Ser Leu Phe Ser Val Leu Val Arg Gly Ser Arg Ala Val
65      145         150         155         160
66      Leu Gln Leu Tyr Pro Glu Asn Ser Glu Gln Leu Glu Leu Ile Thr Thr
67      165         170         175
68      Gln Ala Thr Lys Ala Gly Phe Ser Gly Gly Met Val Val Asp Tyr Pro
69      180         185         190
70      Asn Ser Ala Lys Ala Lys Lys Phe Tyr Leu Cys Leu Phe Ser Gly Pro
71      195         200         205
72      Ser Thr Phe Ile Pro Glu Gly Leu Ser Glu Asn Gln Asp Glu Val Glu
73      210         215         220
74      Pro Arg Glu Ser Val Phe Thr Asn Glu Arg Phe Pro Leu Arg Met Ser
75      225         230         235         240
76      Arg Arg Gly Met Val Arg Lys Ser Arg Ala Trp Val Leu Glu Lys Lys
77      245         250         255
78      Glu Arg His Arg Arg Gln Gly Arg Glu Val Arg Pro Asp Thr Gln Tyr
79      260         265         270
80      Thr Gly Arg Lys Arg Lys Pro Arg Phe
81      275         280
83      (2) INFORMATION FOR SEQ ID NO: 2:
84      (i) SEQUENCE CHARACTERISTICS:
85          (A) LENGTH: 1135 base pairs
86          (B) TYPE: nucleic acid
87          (C) STRANDEDNESS: single
88          (D) TOPOLOGY: linear
89      (vii) IMMEDIATE SOURCE:
90          (A) LIBRARY: THP1PLB01
91          (B) CLONE: 10625
92      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
  
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/024,933

DATE: 01/17/2002
TIME: 17:50:40

Input Set : N:\Crif3\RULE60\10024933.raw
Output Set: N:\CRF3\01172002\J024933.raw

```

93      AGTCGCAGGT GTGCTGCTGA GGCCTGAGAA TGGCGTCCCG CGGCCGGCGT CCGGAGCATG      60
94      GCGGACCCCC AGAGCTGTTT TATGACGAGA CAGAAGCCCG GAAATACTTT CGCAACTCAC      120
95      GGATGATTGA TATCCAGACC AGGATGGCTG GGCAGGCATT GGAGCTTCTT TATCTGCCAG      180
96      AGAATAAGCC CTGTTACCTG CTGGATATTG GCTGTGGCAC TGGGCTGAGT GGAAGTTATC      240
97      TGTCAAGTGA AGGGCACTAT TGGGTGGGCC TGGATATCAG CCGTGCATG CTGGATGAGG      300
98      CTGTGGACCG AGAGATAGAG GGAGACCTGC TGCTGGGGGA TATGGGCCAG GGCATCCCAT      360
99      TCAAGCCAGG CACATTTGAT GGTTCATCA GCATTCTGC TGTGCAGTGG CTCTGTAATG      420
100     CTAACAAGAA GTCTGAAAAC CCTGCCAAGC GCCTGTACTG CTTTTTGTCT TCTCTTTTTT      480
101     CTGTTCTCGT CCGGGGATCC CGAGCTGTCC TGCAGCTGTA CCCTGAGAAC TCAGAGCAGT      540
102     TGGAGCTGAT CACAACCCAG GCCACAAAGG CAGGCTTCTC CGGTGGCATG GTGGTAGACT      600
103     ACCCTAACAG TGCCAAAGCA AAGAAATTCT ACCTCTGCTT GTTTTCTGGG CCTTCGACCT      660
104     TTATACCAGA GGGGCTGAGT GAAAATCAGG ATGAAGTTGA ACCCAGGGAG TCTGTGTTCA      720
105     CCAATGAGAG GTTCCCATTA AGGATGTGCA GGCGGGGAAT GGTGAGGAAG AGTCGGGCAT      780
106     GGGTCTGGA GAAGAAGGAG CGGCACAGGC GCCAGGGCAG GGAAGTCAGA CCGACACCC      840
107     AGTACACCG CCGCAAGCGC AAGCCCCGCT TCTAAGTCAC CACGCGGTTC TGGAAAGGCA      900
108     CTTGCTCTG CACTTTTCTA TATTGTTTCA CTGACAAAGT AGTATTTTAG AAAAGTTCTA      960
109     AAGTTATAAA AATGTTTTCT GCAGTAAAAA AAAAGTTCTC TGGGCCGGGC GTGGTGGCTC      1020
110     ACACCTGTAA TCCAGCACC TTGGGAGGCT GAGGTGGGAG GATCATTGTA GGCCAGGAGT      1080
111     TTGAGACCTG CCGGGCAAC ATAATGAAAC TTCCTTCCCA GGGAGAAAAA AAAAA      1135

```

113 (2) INFORMATION FOR SEQ ID NO: 3:

114 (i) SEQUENCE CHARACTERISTICS:

115 (A) LENGTH: 283 amino acids

116 (B) TYPE: amino acid

117 (C) STRANDEDNESS: single

118 (D) TOPOLOGY: linear

119 (vii) IMMEDIATE SOURCE:

120 (A) LIBRARY: GenBank

121 (B) CLONE: 1065505

122 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

123      Met Ala Ser Phe Lys Val Lys Pro Glu His Thr Gly Pro Pro Asp Leu
124      1          5          10          15
125      Tyr Tyr Asn Glu Thr Glu Ala Ala Lys Tyr Ala Ser Asn Ser His Ile
126      20          25          30
127      Thr Ala Ile Gln His Glu Met Ala Glu Arg Ala Leu Glu Leu Leu Ala
128      35          40          45
129      Leu Pro Glu Gly Lys Ser Gly Phe Leu Leu Asp Ile Gly Cys Gly Thr
130      50          55          60
131      Gly Met Ser Ser Glu Val Ile Leu Asp Ala Gly His Met Phe Val Gly
132      65          70          75          80
133      Val Asp Val Ser Arg Pro Met Leu Glu Ile Ala Arg Gln Asp Glu Asp
134      85          90          95
135      Leu Glu Ser Gly Asp Phe Ile His Gln Asp Met Gly Leu Gly Met Pro
136      100         105         110
137      Phe Arg Pro Gly Ser Phe Asp Gly Ala Ile Ser Ile Ser Ala Ile Gln
138      115         120         125
139      Trp Leu Cys His Ala Asn Ala Ser Asp Glu Asn Pro Arg Lys Arg Leu
140      130         135         140
141      Leu Phe Phe Phe Gln Ser Leu Tyr Gly Cys Leu Gly Arg Gly Ser Arg
142      145         150         155         160

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/024,933

DATE: 01/17/2002

TIME: 17:50:40

Input Set : N:\Crif3\RULE60\10024933.raw

Output Set: N:\CRF3\01172002\J024933.raw

```

143   Ala Val Phe Gln Phe Tyr Pro Glu Asn Asp Glu Gln Cys Asp Leu Ile
144               165                      170                      175
145   Met Gly Gln Ala His Lys Ala Gly Phe Asn Gly Gly Leu Val Val Asp
146               180                      185                      190
147   Phe Pro Glu Ala Ala Lys Arg Lys Lys Val Tyr Leu Val Leu Met Thr
148               195                      200                      205
149   Gly Gly Val Val Gln Leu Pro Gln Ala Leu Thr Glu Asp Gly Glu Glu
150               210                      215                      220
151   Ser Arg Thr Gln Ile Asp Asn Ala Gly Arg Arg Phe Val Trp Asn Ser
152               225                      230                      235                      240
153   Arg Lys Asn Glu Lys Val Ala Lys Gly Ser Lys Ala Trp Ile Glu Ala
154               245                      250                      255
155   Lys Arg Gln Arg Gln Ile Lys Gln Gly Arg Asp Val Arg His Glu Ser
156               260                      265                      270
157   Lys Tyr Ser Gly Arg Lys Arg Lys Thr Lys Phe
158               275                      280

```

160 (2) INFORMATION FOR SEQ ID NO: 4:

161 (i) SEQUENCE CHARACTERISTICS:

162 (A) LENGTH: 275 amino acids

163 (B) TYPE: amino acid

164 (C) STRANDEDNESS: single

165 (D) TOPOLOGY: linear

166 (vii) IMMEDIATE SOURCE:

167 (A) LIBRARY: GenBank

168 (B) CLONE: 1907189

169 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

170   Met Ser Arg Pro Glu Glu Leu Ala Pro Pro Glu Ile Phe Tyr Asn Asp
171       1           5           10           15
172   Ser Glu Ala His Lys Tyr Thr Gly Ser Thr Arg Val Gln His Ile Gln
173               20           25           30
174   Ala Lys Met Thr Leu Arg Ala Leu Glu Leu Leu Asn Leu Gln Pro Cys
175               35           40           45
176   Ser Phe Ile Leu Asp Ile Gly Cys Gly Ser Gly Leu Ser Gly Glu Ile
177       50           55           60
178   Leu Thr Gln Glu Gly Asp His Val Trp Cys Gly Leu Asp Ile Ser Pro
179       65           70           75           80
180   Ser Met Leu Ala Thr Gly Leu Ser Arg Glu Leu Glu Gly Asp Leu Met
181               85           90           95
182   Leu Gln Asp Met Gly Thr Gly Ile Pro Phe Arg Ala Gly Ser Phe Asp
183               100          105          110
184   Ala Ala Ile Ser Ile Ser Ala Ile Gln Trp Leu Cys Asn Ala Asp Thr
185               115          120          125
186   Ser Tyr Asn Asp Pro Lys Gln Arg Leu Met Arg Phe Phe Asn Thr Leu
187       130          135          140
188   Tyr Ala Ala Leu Lys Lys Gly Gly Lys Phe Val Ala Gln Phe Tyr Pro
189       145          150          155          160
190   Lys Asn Asp Asp Gln Val Asp Asp Ile Leu Gln Ser Ala Lys Val Ala
191               165          170          175
192   Gly Phe Ser Gly Gly Leu Val Val Asp Asp Pro Glu Ser Lys Lys Asn

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/024,933

DATE: 01/17/2002

TIME: 17:50:40

Input Set : N:\Crf3\RULE60\10024933.raw

Output Set: N:\CRF3\01172002\J024933.raw

193		180		185		190													
194	Lys	Lys	Tyr	Tyr	Leu	Val	Leu	Ser	Ser	Gly	Ala	Pro	Pro	Gln	Gly	Glu			
195			195					200					205						
196	Glu	Gln	Val	Asn	Leu	Asp	Gly	Val	Thr	Met	Asp	Glu	Glu	Asn	Val	Asn			
197		210					215					220							
198	Leu	Lys	Lys	Gln	Leu	Arg	Gln	Arg	Leu	Lys	Gly	Gly	Lys	Asp	Lys	Glu			
199	225				230					235						240			
200	Ser	Ala	Lys	Ser	Phe	Ile	Leu	Arg	Lys	Lys	Glu	Leu	Met	Lys	Arg	Arg			
201				245						250					255				
202	Gly	Arg	Lys	Val	Ala	Lys	Asp	Ser	Lys	Phe	Thr	Gly	Arg	Lys	Arg	Arg			
203			260						265					270					
204	His	Arg	Phe																
205			275																

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/024,933

DATE: 01/17/2002

TIME: 17:50:41

Input Set : N:\CrF3\RULE60\10024933.raw

Output Set: N:\CRF3\01172002\J024933.raw

L:22 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:23 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]